

Latency-Energy Tradeoff with Realistic Hardware Models

A. Høst-Madsen, N. Whitcomb, J. Weldon
 Department of Electrical Engineering
 University of Hawaii, Manoa
 Honolulu, HI, 96822
 Email: {ahm,nwhitcom,jaweldon}@hawaii.edu

Zixiang Xiong
 Department of Electrical and Computer Engineering
 Texas A&M University,
 College Station, TX 77843
 E-mail: zx@ece.tamu.edu

Abstract—Latency and energy are two fundamental parameters that characterize communications systems performance. In this paper we investigate the fundamental relationship between these quantities using finite blocklength information theory. We take into account realistic models of hardware, including overhead power, power amplifier inefficiency and the receiver noise factor. With these models we find that energy and latency behaves quite different from the ideal case.

I. INTRODUCTION

The focus of this paper is to understand the relationship between delay and energy in wireless communications, taking into consideration realistic models of communications hardware. With the proliferation of mobile devices, such as smart phones and tablet PCs, wireless communications are increasingly used to serve traffic with stringent delay constraints, such as video streaming, online gaming, VoIP, and video conferencing. Nowadays most of the internet traffic are video. Therefore, providing stringent delay guarantees becomes an important challenge for enhancing the quality of service (QoS) of end users.

On the other hand, energy consumption of communications is becoming an increasing focus under the banner of “green” communications (IEEEExplore returns thousands of hits on “green communications”). The main reason for this is the prevalence of mobile battery powered devices. The two trends, delay sensitive communications and the desire for energy conservation, are basically conflicted. Hence it is important to understand the basic tradeoff between energy and delay. In [1] we analyzed this relationship for idealized models of communications hardware. The current paper will extend this analysis to more realistic hardware models.

II. STREAMING WITH LATENCY CONSTRAINTS

We consider the following streaming model. Consider an AWGN (additive white Gaussian noise) channel with symbol spacing T_c . An infinite stream $b[t]$, $t = 1, 2, \dots$ of bits arrive periodically at a transmitter with spacing T_s , i.e., arrival rate $\lambda = T_s^{-1}$; we let $R_a = \frac{T_c}{T_s} = \lambda T_c$ be the unit-less arrival rate. Equivalently, the bandwidth for transmission is $B = T_c^{-1} = R_a^{-1} T_s^{-1}$. The decoder needs to decode bit $b[t]$ no later than at time $(t + d)T_s$, where d is the (unit-less) delay.

Supported in part by NSF grants EECS-1923751 and EECS-1923803.

For a finite delay and finite energy, error-free transmission is not possible. The most natural setting for the problem we consider is clearly sequential decoding, which was studied by Fano in [3] with practical coding considered in [5]. However, in this work we will only consider block coding. There are a number of reasons for this: packet based transmission is used in most practical communications systems, practical block coding is more developed than sequential coding, and we can use the theory initiated with [7].

For packet transmission, the transmitter takes k bits from the input bit stream and packs them into a packet. This packet is then transmitted in n channel uses; all the bits must arrive at the receiver within dT_s seconds after the first bit in the packet arrived at the transmitter, see Fig. 1 (we assume zero transmission and decoding delay). We assume that a packet is either received without error or is lost with probability δ . This means that independent of the packet length k , the fraction of bits lost is δ . We consider δ as a fixed and given constraint in our system independent of R_a and d .

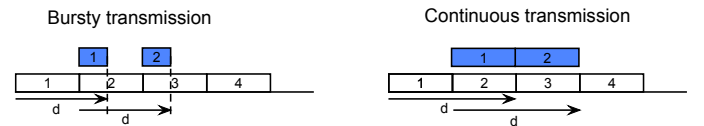


Fig. 1. Packet transmission modes. Blue are transmitted packets.

From Fig. 1 we get the following constraints

$$\frac{d}{2} \leq k \leq d; \quad n = \frac{d - k}{R_a} \quad (1)$$

The energy to transmit the packet is

$$\frac{E_b}{\frac{N_0}{2}} = \frac{nP}{k} \quad (2)$$

For simplicity we set $N_0 = 1$ so that

$$E_b = \frac{nP}{2k}. \quad (3)$$

with P the transmit power.

From [7], [8] we get the following relationship between n and k ,

$$k = nC(P) - \sqrt{nV(P)}Q^{-1}(\epsilon) + \frac{1}{2} \log n + O(1). \quad (4)$$

Here $Q(\cdot)$ is the Q-function,

$$V(P) = \frac{P}{2} \frac{P+2}{(P+1)^2} \log^2 e \quad (5)$$

is the channel dispersion, and

$$C(P) = \frac{1}{2} \log(1+P) \quad (6)$$

is the channel capacity. Previously, in order to use information theory to analyze delay, the packet size k would have to converge to infinity to use asymptotic results. This is of course a contradiction: with infinite packet size k , the delay d is infinite. That can be circumvented somewhat by carefully scaling different time scales jointly and suitable interpretation. Still, it leads to somewhat convoluted results. With finite blocklength theory it we are finally able to analyze, for example, what is the actual energy needed to meet a given delay constraint.

In [1] we derived rigorous results for this problem. When we fix R_a and let $d \rightarrow \infty$, we have the following result:

Theorem 1. For fixed R_a the energy per bit is given by

$$E_b(d) = \frac{P_0}{2R_a} + \frac{(P_0+1)\sqrt{2V_0P_0}}{\sqrt{R_a} \log e} Q^{-1}(\delta) \sqrt{d^{-1}} + o(\sqrt{d^{-1}}) \quad (7)$$

where

$$V_0 = \frac{P_0(P_0+2)}{2(P_0+1)^2} \log^2 e$$

and $P_0 = 2^{2R_a} - 1$. This limit can be achieved by continuous transmission.

III. HARDWARE CONSTRAINTS

In [1] we have assumed ideal models of hardware for communications. The aim of this paper is to include more realistic models of communications hardware. This can quite dramatically change some conclusions. For example, in the ideal model time and bandwidth are equivalent, i.e., wideband communications is equivalent to slow, narrowband communications with respect to energy. However, from a hardware perspective these are no longer equivalent. As an example, if there is a cost of staying on, slow communications might not be energy efficient. How hardware constraints influence energy from an information theory point of view has been considered before. However, as mentioned above, when there is a delay constraint, one has to consider finite blocklength, and more realistic insights to the relationship between energy and hardware can therefore be obtained by using finite blocklength theory.

A. A Simple Hardware Model

In this section, we will show how the results in [1] can be extended to take into account hardware constraints. In the ideal model, the capacity of an AWGN channel is given by (6), which we rewrite as

$$C = \frac{1}{2} \log(1 + \tilde{P}) \quad (8)$$

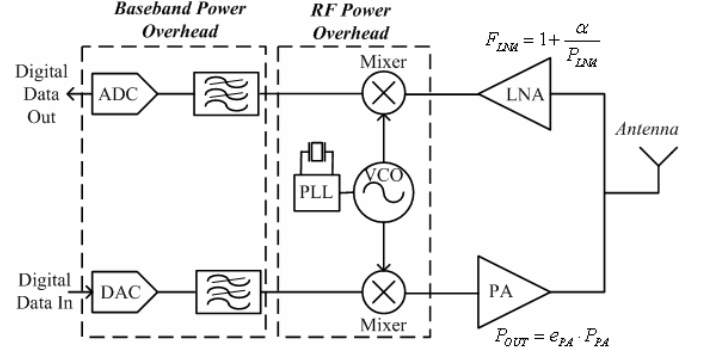


Fig. 2. Hardware model, from [2]. This paper uses e_{LNA} for α in the figure.

where $\tilde{P}$ now denotes the received SNR, with E_b given by

$$E_b = \frac{n\tilde{P}}{2k} \quad (9)$$

The thesis [2] considered the interplay between information theory and hardware design in detail. We will consider the models in [2], see Fig. 2. According to [2, Equation (3.2)] the power consumed by the transmitter is

$$P_{TX} = P_{OH-TX} + \frac{1}{e_{PA}} P_{OUT}.$$

Here P_{OUT} is the power that can actually be used for transmission, e_{PA} is power amplifier efficiency factor, and P_{OH-TX} is an overhead power. The power consumed by the receiver is $P_{RX} = P_{OH-RX} + P_{LNA}$ where P_{LNA} is the power consumed by the low noise amplifier (LNA). The LNA affects the system performance through the receiver noise factor [2, Equation (3.8)]¹ $F_{LNA} = 1 + \frac{e_{LNA}}{P_{LNA}}$.

Based on this model there are many different objectives and models that can be considered. We consider the case where the desire is to minimize total energy consumption and the transceivers are able to completely switch off when not transmitting. If we desire to find the total energy per bit consumed by the system (by both transmitter and receiver), we can take this into account as follows. We replace (8) by

$$C = \frac{1}{2} \log(1 + P) = C(P) \quad (10)$$

where

$$P = \gamma \frac{P_{OUT}}{F_{LNA}}, \quad (11)$$

with γ the path loss in the channel, is what we could call the equivalent received power, and we replace (9) by

$$E_b = \frac{n(P_{TX} + P_{RX})}{2k} \quad (12)$$

The objective is to minimize the total energy consumption by adjusting P_{OUT} and P_{LNA} . We solve this as follows. Suppose

¹We use e_{LNA} for α in [2] as α has a different meaning below.

P is given; the objective is to minimize $\frac{P_{OUT}}{e_{PA}} + P_{LNA}$. It is easily seen that the solution is

$$P_{LNA} = \sqrt{\frac{e_{LNA}P/\gamma}{e_{PA}}} \\ P_{OUT} = P + \sqrt{e_{LNA}e_{PA}P/\gamma} \quad (13)$$

We first consider what happens in the limit as $d \rightarrow \infty$. Let α be the duty cycle. Then $k = (1 - \alpha)d$, $n = d\alpha R_a^{-1}$, where $\alpha \in (0, \frac{1}{2}]$ due to (1). In the limit $d \rightarrow \infty$

$$\frac{k}{n} = \frac{1 - \alpha}{\alpha} R_a = C(P)$$

or $P(\alpha) = C^{-1}(\frac{1 - \alpha}{\alpha} R_a) = 2^2(\frac{1 - \alpha}{\alpha} R_a) - 1$ and

$$E_b = \frac{n P_{TOT}}{2k} = \frac{\alpha (P_{OH} + e_{PA}^{-1} P_{OUT} + P_{LNA})}{2(1 - \alpha) R_a} \\ = \frac{\alpha \left(P_{OH} + \frac{P(\alpha)}{\gamma e_{PA}} + 2\sqrt{\frac{e_{LNA}P(\alpha)}{\gamma e_{PA}}} \right)}{2(1 - \alpha) R_a} \quad (14)$$

with $P_{OH} = P_{OH-TX} + P_{OH-RX}$. Notice that the energy depends only on the product γe_{PA} , and in the following we there include γ in e_{PA} . We can now minimize E_b over α , which must be done numerically.

Denote by α_0 and P_0 the optimum solution to (14) with C_0 the corresponding value in (10).

Since $\alpha \leq \frac{1}{2}$, for some values of the parameters we may have $\alpha_0 = \frac{1}{2}$ and then $C_0 = R_a$ and $P_0 = 2^{2R_a} - 1$ corresponding to continuous transmission, while for others $\alpha_0 < \frac{1}{2}$, $C_0 > R_a$ and $P_0 > 2^{2R_a} - 1$, bursty transmission. As opposed to the ideal solution, bursty transmission might be more energy efficient than continuous transmission. This has been observed previously in [4] in a slightly different context.

If we keep all other parameters fixed and decrease P_{OH} , α_0 will increase until it hits $\frac{1}{2}$ for some value $P_{OH}^* > 0$. Thus, for $P_{OH} \leq P_{OH}^*$, continuous transmission minimizes energy. We can explicitly calculate P_{OH}^* by solving $\frac{\partial E_b}{\partial \alpha} \Big|_{\alpha=\frac{1}{2}} < 0$ getting

$$P_{OH}^* = \frac{2(P_c + 1)R_a \ln 2 - P_c}{e_{PA}} \\ + 2 \frac{(P_c + 1)R_a \ln 2 - P_c}{e_{PA}} \sqrt{\frac{e_{LNA}}{e_{PA}P_c}}$$

where $P_c = 2^{2R_a} - 1$ is the power required for continuous transmission. It is possible that $P_{OH}^* < 0$, meaning that even without overhead power, bursty transmission is optimum. This happens if

$$e_{LNA} > \frac{P_c}{4e_{PA}} \frac{(2(P_c + 1)R_a \ln 2 - P_c)^2}{((P_c + 1)R_a \ln 2 - P_c)^2}$$

B. Finite Blocklength

We can now generalize Theorem 1 to take into account the more realistic hardware model above. The result does depend on α_0 and P_0 , but given those we can get closed form solutions, as follows

Theorem 2. If $\alpha_0 = \frac{1}{2}$,

$$E_b(d) = \frac{P_{TOT}}{2R_a} + \beta \frac{(P_0 + 1)\sqrt{2V_0P_0}}{\sqrt{R_a} \log e} Q^{-1}(\delta) \sqrt{d^{-1}} \\ + o(\sqrt{d^{-1}}) \quad (15)$$

where $\beta = e_{PA}^{-1} + \sqrt{\frac{e_{LNA}}{e_{PA}P_0}}$ and $P_{TOT} = P_{OH} + \frac{P_0}{\gamma e_{PA}} + 2\sqrt{\frac{e_{LNA}P_0}{\gamma e_{PA}}}$. On the other hand, for $\alpha_0 < \frac{1}{2}$ we get

$$E_b(d) = \frac{\alpha_0 P_{TOT}}{2(1 - \alpha_0)R_a} \\ + \sqrt{\frac{V_0}{(C_0 + R_a)^3}} \frac{P_{TOT}}{2(\alpha_0 - 1)^2} Q^{-1}(\delta) \sqrt{d^{-1}} + o(\sqrt{d^{-1}}) \quad (16)$$

This can be achieved by letting

$$P = P_0 + K\sqrt{d^{-1}} \\ \alpha = \alpha_0 + \left(R_a \sqrt{\frac{V_0}{(C_0 + R_a)^3}} Q^{-1}(\delta) \right. \\ \left. - K \frac{R_a}{2 \ln 2 (P_0 + 1)(C_0 + R_a)^2} \right) \sqrt{d^{-1}} \quad (17)$$

where $K \in \mathbb{R}$ is arbitrary.

Proof: The case of $\alpha_0 = \frac{1}{2}$ is a small variation to [1], and we will not repeat the proof here. We will therefore consider the case $\alpha_0 < \frac{1}{2}$. We can write (4) in terms of α ,

$$(1 - \alpha) = \alpha R_a^{-1} C - \sqrt{d^{-1}} \alpha R_a^{-1} V Q^{-1}(\delta) \\ + \frac{1}{2} \frac{\log(\alpha d R_a^{-1})}{d} + \frac{b(n, P)}{d}, \quad (18)$$

In this case, both P and α are variable. First we solve (18) with respect to α , up to terms of order $o(\sqrt{d^{-1}})$. To reduce equation size, in the following let $\tilde{V} = V(Q^{-1}(\delta))^2$

$$\alpha = \frac{\sqrt{-4C d o(\sqrt{d^{-1}}) R_a^2 \tilde{V} + 4C d R_a^2 \tilde{V} - 4d o(\sqrt{d^{-1}}) R_a^3 \tilde{V} + 4d R_a^3 \tilde{V} + R_a^2 \tilde{V}^2}}{2(C^2 + 2C R_a + R_a^2)} \\ = \frac{2C o(\sqrt{d^{-1}}) R_a + 2C R_a + \frac{R_a \tilde{V}}{d} - 2o(\sqrt{d^{-1}}) R_a^2 + 2R_a^2}{2(C^2 + 2C R_a + R_a^2)} \\ = \frac{\sqrt{4C d R_a^2 \tilde{V} + 4d R_a^3 \tilde{V}}}{2(C^2 + 2C R_a + R_a^2)} + o(\sqrt{d^{-1}}) \\ = \frac{2R_a \sqrt{(C + R_a) \tilde{V}} \sqrt{d^{-1}} + 2C R_a + 2R_a^2}{2(C^2 + 2C R_a + R_a^2)} + o(\sqrt{d^{-1}}) \\ = \frac{2C R_a + 2R_a^2}{2(C^2 + 2C R_a + R_a^2)} + \frac{2R_a \sqrt{(C + R_a) \tilde{V}}}{2(C^2 + 2C R_a + R_a^2)} \sqrt{d^{-1}} \\ + o(\sqrt{d^{-1}}) \\ = R_a \frac{1}{C + R_a} + R_a \sqrt{\frac{V}{(C + R_a)^3}} Q^{-1}(\delta) \sqrt{d^{-1}} + o(\sqrt{d^{-1}}) \quad (19)$$

We can expand this as a series in ΔP as follows

$$\alpha = \alpha_0 + A_0\sqrt{d^{-1}} + A_1\Delta P + (A_2\Delta P + o(\Delta P))\sqrt{d^{-1}} + o(\Delta P) + o(\sqrt{d^{-1}}) \quad (20)$$

and find

$$\begin{aligned} \alpha_0 &= \frac{R_a}{C_0 + R_a} \\ A_0 &= R_a \sqrt{\frac{V_0}{(C_0 + R_a)^3}} Q^{-1}(\delta) \\ A_1 &= -\frac{R_a}{2\ln(2)(P_0 + 1)(C_0 + R_a)^2} \end{aligned}$$

Equation (14) is still true, but now with $P(\alpha)$ given for finite d , specifically from (19), as follows

$$E_b = \frac{\alpha(P) \left(P_{OH} + \frac{P}{\gamma e_{PA}} + 2\sqrt{\frac{e_{LNA}P}{\gamma e_{PA}}} \right)}{2(1 - \alpha(P))R_a} \quad (21)$$

Here

$$\begin{aligned} \frac{\alpha(P)}{1 - \alpha(P)} &= \frac{\alpha_0}{1 - \alpha_0} + \frac{\Delta\alpha}{(1 - \alpha_0)^2} + \frac{\Delta\alpha^2}{(1 - \alpha_0)^3} + o(\Delta\alpha^2) \\ &= \frac{\alpha_0}{1 - \alpha_0} \\ &\quad + \frac{1}{(1 - \alpha_0)^2} \left(A_0\sqrt{d^{-1}} + A_1\Delta P \right) \\ &\quad + \frac{1}{(1 - \alpha_0)^2} \left((A_2\Delta P + o(\Delta P))\sqrt{d^{-1}} + o(\Delta P) \right) \\ &\quad + \frac{1}{(1 - \alpha_0)^3} \left(A_0\sqrt{d^{-1}}A_1\Delta P \right) \\ &\quad + o(\sqrt{d^{-1}}) \end{aligned} \quad (22)$$

and

$$\begin{aligned} P_{OH} + \frac{P}{\gamma e_{PA}} + 2\sqrt{\frac{e_{LNA}P}{\gamma e_{PA}}} \\ &= P_{OH} + \frac{P_0}{\gamma e_{PA}} + 2\sqrt{\frac{e_{LNA}P_0}{\gamma e_{PA}}} \\ &\quad + \frac{\Delta P}{\gamma e_{PA}} + \sqrt{\frac{e_{LNA}P_0}{\gamma e_{PA}}} \frac{\Delta P}{P_0} + o(\Delta P) \end{aligned} \quad (23)$$

Let $f(d) = \Delta P = P - P_0$. We must have $\lim_{d \rightarrow \infty} f(d) = 0$. It is also seen from (22) and (23) that if $f(d)$ increases at a rate faster than $\sqrt{d^{-1}}$, this will lead to E_b also increasing at a faster rate, while any order lower than $\sqrt{d^{-1}}$ can be included in the $o(\sqrt{d^{-1}})$ term. We can therefore put

$$\begin{aligned} \Delta P &= C_2\sqrt{d^{-1}} \\ \Delta\alpha &= C_1\sqrt{d^{-1}} = A_0\sqrt{d^{-1}} + A_1C_2\sqrt{d^{-1}} \end{aligned} \quad (24)$$

We expand (21) to first order in $\sqrt{d^{-1}}$ using (24)

$$\begin{aligned} E_b &= E_{b,\min} + \frac{\left(\alpha_0^2(-C_2) \left(e_{PA}\sqrt{\frac{e_{LNA}P_0}{e_{PA}}} + P_0 \right) \right)}{2(\alpha_0 - 1)^2 e_{PA} P_0 R_a} \sqrt{d^{-1}} \\ &\quad + \frac{\left(\alpha_0 C_2 \left(e_{PA}\sqrt{\frac{e_{LNA}P_0}{e_{PA}}} + P_0 \right) \right)}{2(\alpha_0 - 1)^2 e_{PA} P_0 R_a} \sqrt{d^{-1}} \\ &\quad + \frac{\left(C_1 P_0 \left(e_{PA} \left(2\sqrt{\frac{e_{LNA}P_0}{e_{PA}}} + P_{OH} \right) + P_0 \right) \right)}{2(\alpha_0 - 1)^2 e_{PA} P_0 R_a} \sqrt{d^{-1}} \end{aligned}$$

which for $C_2 = 0$ gives

$$\begin{aligned} E_b &= E_{b,\min} + \frac{A_0 \left(e_{PA} \left(2\sqrt{\frac{e_{LNA}P_0}{e_{PA}}} + P_{OH} \right) + P_0 \right)}{2(\alpha_0 - 1)^2 e_{PA} R_a} \sqrt{d^{-1}} \\ &= E_{b,\min} \\ &\quad + \sqrt{\frac{V_0}{(C_0 + R_a)^3}} \frac{e_{PA} \left(2\sqrt{\frac{e_{LNA}P_0}{e_{PA}}} + P_{OH} \right) + P_0}{2(\alpha_0 - 1)^2 e_{PA}} \\ &\quad \times Q^{-1}(\delta) \sqrt{d^{-1}} \end{aligned}$$

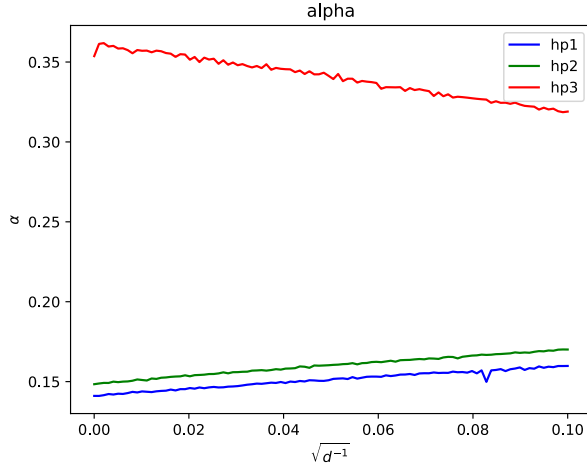
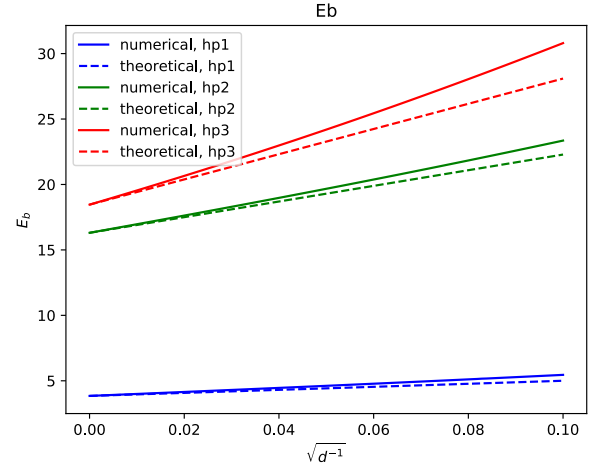
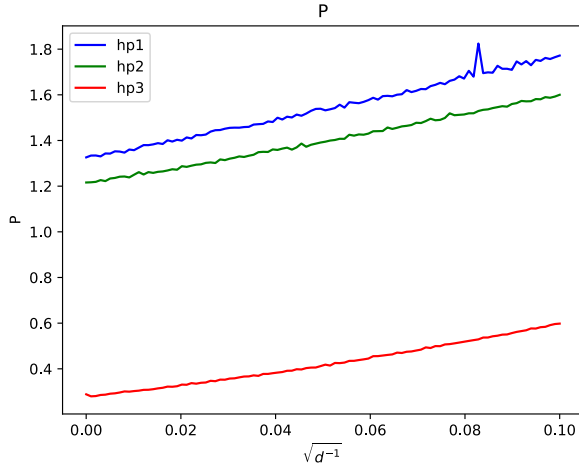
This result might be somewhat surprising, specifically (19). Consider the following question. How should we compensate for decreasing d ? In Theorem 1 and in (15) this is done by keeping α fixed at $\frac{1}{2}$ (continuous transmission) and increasing the transmission power P . When $\alpha_0 < \frac{1}{2}$, resulting in bursty transmission for infinite delay, one might ask if it is optimum to increase the duty cycle or the power to compensate for decreasing delay. But (19) shows that this question has no answer. Since K can be both positive and negative, one could increase or decrease P , which might lead to increase or decrease in α . The resulting increase in E_b is always the same, as long as (19) is observed. This seems paradoxical: couldn't one then change α_0 and P_0 arbitrarily, contradicting the claim that α_0 and P_0 is an optimum solution? But exactly because this is a minimum, changing α and P does not change E_b to the first order. On the other hand, changes in α and P to compensate for decreasing d only concerns first order changes. While this seems odd, it is confirmed by numerical results. ■

IV. NUMERICAL RESULTS

In this section, we compare the theoretical results with "simulated" results. Specifically, we put $b(n, P) = 0$ in (20) (i.e., ignore the $O(1)$ term in (4)) and solve numerically for P as a function of α . Then we insert that in (14) and optimize numerically over α . We use the following parameters

$$\begin{aligned} R_a &= 1 \\ \delta &= 0.001 \\ P_{OH} &\in \{1, 5, 0.5\} \\ e_{LNA} &\in \{1, 0.05, 0.01\} \\ e_{PA} &\in \{0.5, 0.1, 0.05\} \end{aligned}$$

(named hp1, hp2, hp3). The results can be seen in Fig. 3-5. The optimization is well-posed and with a distinct minimum

Fig. 3. Duty cycle α as a function of d .Fig. 5. Energy per bit E_b as a function of d .Fig. 4. Power P as a function of d .

and gives specific functions for P and α as a function of d . Yet, the resulting E_b agrees with the theory, where P and α are arbitrary. Are the optimum functions illusions? The optimization does take into account more terms, so perhaps they are indeed optimum, but the theory shows that changing them will have essentially no influence on E_b , as long as (19) is satisfied.

V. CONCLUSION

In this paper we have shown how to extend the finite delay theory in [1] to more realistic hardware models. Now, it is well known that to approach the Shannon minimum energy per bit

$$(E_b/N_0)_{\min} = \ln 2 = -1.59 \text{ dB}, \quad (25)$$

it is necessary to have $P \rightarrow 0$, or in our model, $R_a \rightarrow 0$. In [6], [1] this limit was considered for ideal hardware models. The interesting question is how this limit behaves for more realistic hardware models; probably one will have to look more fundamentally at hardware behavior than the models in [2].

REFERENCES

- [1] Mirza Uzair Baig, Lei Yu, Zixiang Xiong, Anders Høst-Madsen, Houqiang Li, and Weiping Li. On the energy-delay tradeoff in streaming data: Finite blocklength analysis. *IEEE Transactions on Information Theory*, 66(3):1861–1881, March 2020.
- [2] Ben Walter Cook. *Low Energy RF Transceiver Design*. PhD thesis, EECS, University of California at Berkeley, 2007.
- [3] R. Fano. A heuristic discussion of probabilistic decoding. *Information Theory, IEEE Transactions on*, 9(2):64–74, 1963.
- [4] P.M. Maassaad, M. Medard, and L. Zheng. Impact of processing energy on the capacity of wireless channels. In *International Symposium on Information Theory and its Applications. ISITA04*, Oct. 2004.
- [5] S.V. Maiya, D.J. Costello, and T.E. Fuja. Low latency coding: Convolutional codes vs. ldpc codes. *Communications, IEEE Transactions on*, 60(5):1215–1225, 2012.
- [6] Y. Polyanskiy, H. V. Poor, and S. Verdú. Feedback in the non-asymptotic regime. *IEEE Transactions on Information Theory*, 57(8):4903–4925, Aug 2011.
- [7] Y. Polyanskiy, H.V. Poor, and S. Verdú. Channel coding rate in the finite blocklength regime. *Information Theory, IEEE Transactions on*, 56(5):2307–2359, may 2010.
- [8] V. Y. F. Tan and M. Tomamichel. The third-order term in the normal approximation for the AWGN channel. *IEEE Transactions on Information Theory*, 61(5):2430–2438, May 2015.